Oracle Data Guard Deep Dive

Emre Baransel
Turkcell
Ankara - Turkey

Keywords:
Data Guard, Performance Tuning, Switchover, Failover, Corruption Detection, integration

Introduction

This session covers functional details and best practices about the implementation and administration of Oracle Data Guard, the replication and disaster recovery solution for Oracle Databases. Data Guard performance tuning, role change best practices, integration issues will be addressed in this presentation.

Configuration Considerations

Choosing the Protection Mode

MODE	REDO TRANSPORT	ACTION WITH NO STANDBY DATABASE CONNECTION	RISK OF DATA LOSS
Maximum Protection	SYNC & LGWR	The primary database needs to write redo to at least one standby database. Otherwise it will shut down.	Zero data loss is guaranteed.
Maximum Availability	SYNC & LGWR	Normally works with SYNC redo transport. If primary database cannot write redo to any of its standby databases, it continues processing transactions as in ASYNC mode.	Zero data loss in normal operation, but not guaranteed
Maximum Performance	ASYNC & (LGWR or ARCH)	Never expects acknowledgment from standby database.	Potential for minimal data loss in normal operation

When any data loss is not acceptable, make your network bandwidth high enough and use Maximum Protection. (Service outage is preferred against any data loss)

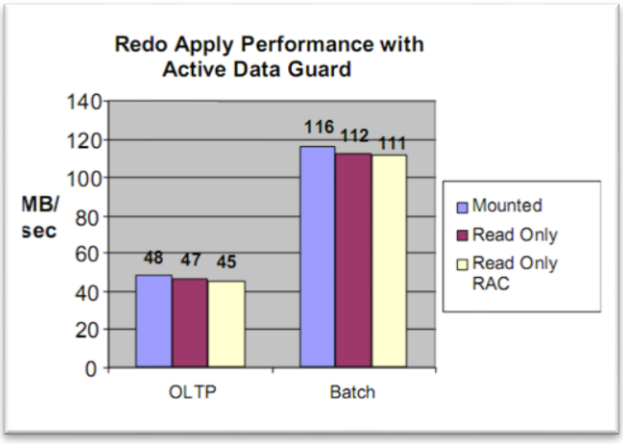
If there is no intolerance about data loss and have high bandwidth use Maximum Availability. If SYNC redo transport is chosen in an 11g Data Guard configuration, the performance decrease on the primary database will be less than the earlier releases. Previously, primary database was first finishing writes to online redo log and then sending redo to standby database. There were two consecutive I/O operations that primary database needs to wait in order to complete the commit. In 11g these two I/O operations run in parallel. Primary database does not wait finishing writes to online redo log and it sends the redo data to standby at the same time.

If there is network latency issue then use Maximum Performance (ASYNC) with LGWR. ARCH is not recommended because it has not any performance benefit but has less data protection in 11g. With LGWR increase log buffer size if necessary, this keeps NSA process reading from memory.

Prefer Real Time Apply with “Flashback On” rather than “Delay”. Delayed configuration increases RTO.

Performance Tuning

Redo Apply Tuning:



Oracle test results reach 48MB/s apply rate for OLTP and 116MB/s for OLAP type workload. Find the redo apply rate on your standby to assess.

1. Method:

In the first method v$recovery_progress view shows two values Active and avarage apply rate. Active rate may be very nonstable where avarage rate gives more accurate idea. You can use this view if you're not using real-time apply because it counts the time that standby waits for the new redo to arrive, and it shows less then the actual value. So it doesn't provide accurate results with real-time apply.

```
SQL> select * from v$recovery_progress
23-SEP-11 Media Recovery   Active Apply Rate   KB/sec  15564   0
23-SEP-11 Media Recovery   Average Apply Rate  KB/sec  20890   0
```

2. Method:

Second method uses V$STANDBY_APPLY_SNAPSHOT view. This view shows reliable values even you use real time apply, whereas this view was deprecated in 11gR2. If you're on a relase under 11gR2 this view is useful to determine the apply rate.

```
SQL> select APPLY_RATE  from V$STANDBY_APPLY_SNAPSHOT;
APPLY_RATE
----------
16305
```

3. Method:

And the last method needs some calculation. It's the method that is recommended on Oracle's Data Guard Best Practices guide. First you find the redo block size with the first query. Then with the second query you monitor the MRP process and check which block of the archivelog is being applied, then after some seconds you query again. Then using the formulation you calculate the redo apply rate.

```
SQL> SELECT lebsz LOG_BLOCK_SIZE from x$kccle; → Redo block size(byte)
SQL>  SELECT PROCESS, SEQUENCE#, THREAD#, block#, BLOCKS,
TO_CHAR(SYSDATE, 'DD-MON-YYYY HH:MI:SS') time from v$MANAGED_STANDBY
WHERE PROCESS='MRP0';
```

```
PROCESS    SEQUENCE#  THREAD#  BLOCK#   BLOCKS     TIME
---------  ---------- ----     -------  --------   --------------
MRP0       276877     1        147338   4097947    19-APR-2012 12:25:34

PROCESS    SEQUENCE#  THREAD#  BLOCK#   BLOCKS     TIME
---------  ---------- ----     -------  --------   --------------------
MRP0       276877       1      645542   4097947    19-APR-2012 12:25:39
```

Media Recovery Rate:

((BLOCK#_END – BLOCK#_BEG) * LOG_BLOCK_SIZE)) / ((TIME_END – TIME_BEG) * 1024 * 1024)

In order to tune Redo Apply rate:

- By default recovery parallelism equals to CPU Count-1. Do not use any other values.
- Keep PARALLEL_EXECUTION_MESSAGE_SIZE >= 8192
- Keep DB_CACHE_SIZE >= Primary value
- Keep DB_BLOCK_CHECKING = FALSE (if you have to)
- System Resources Needs to be checked
- Query what MRP process is waiting

```
SQL> select a.sid, b.username, b.osuser, a.event, a.wait_time, a.p1, a.p1text,
a.seconds_in_wait from gv$session_wait a, gv$session b where a.sid=b.sid and
b.sid=(select SID from v$session where PADDR=(select PADDR from v$bgprocess
where NAME='MRP0'));
```

Redo Transport Tuning

Tune LOG_ARCHIVE_MAX_PROCESS parameter on the primary database. This parameter specifies the parallelism of redo transport between primary and standby servers. Default value is 2 but this will be generally not sufficient for Data Guard configurations. In a configuration with high redo generation rate and multiple standbys this value must be increased up to 30. Redo transport parallelism significantly increases redo transport rate.

Consider using Redo Transport Compression: In Oracle Database 11g Release 2 (11.2.0.2) redo transport compression is no longer limited to compressing redo data only when a redo gap is being resolved. Compression is always on. If the CPU power is available and there is a low bandwidth network, consider using redo compression.

Also consider:

- Configuring TCP Send / Receive Buffer Sizes
- Increasing SDU Size
- Setting TCP.NODELAY to YES

Role Transition Best Practices

Switchover Best Practices:

- Stop job processing by setting the AQ_TM_PROCESSES parameter to 0.
- Configure the standby database to use real-time apply and, if possible, ensure the databases are synchronized before the switchover operation to optimize switchover processing.

- For a physical standby database, reduce the number of archiver (ARCn) processes to the minimum needed for both remote and local archiving.
- Properly set archiving destinations on the Standby database.
- In 11gR2 set "_ktb_debug_flags"=8 for the "Bug 8895202 - ORA-1555 / ORA-600 [ktbdchk1: bad dscn] in Physical Standby after switch-over"

Failover Best Practices:

- Enable Flashback Database to reinstate the failed primary databases after a failover operation has completed. Flashback Database facilitates fast point-in-time recovery, if needed.
- Use real-time apply with Flashback Database to apply redo data to the standby database as soon as it is received, and to quickly rewind the database should user error or logical corruption be detected.
- Consider configuring multiple standby databases to maintain data protection following a failover.

Corruption Detection

First of all the introduction of 'end-to-end checksums' in Oracle Database 11g makes it unnecessary to set DB_BLOCK_CHECKING parameter on the standby database in order to detect corruptions that may occur at the primary database. As i mentioned before in 10g setting this parameter to TRUE has significant effect on recovery performance.This feature does not guarantee to detect all logical corruptions but in our tests we see that it works for most.

The 'Automatic Block Corruption Repair' feature came with 11g R2 Active Data Guard. Oracle documentation says that in order to use this feature you must use Physical Standby & Maximum Availability mode. But in one of our maximum performance mode Data Guard configuration we saw that production database is trying to use this auto-repair feature and raised an SR. Then Oracle said that just Physical Standby is necessary for this feature, it doesn't have to be in Max. Availability mode.

Corruption on the primary can be repaired by using standby as the source, also corruption on the standby can be repaired by using primary as the source.

Also using RMAN "RECOVER BLOCK" command you can repair the corruption. This operation will try use the standby database first. If you don't want to use the standby database for corruption repair, you must use EXCLUDE STANDBY option in the "RECOVER BLOCK" command.

"Lost – Write" detection is also an 11g feature. It came with 11gR1.This is a serious corruption which has its source in I/O subsystem. Storage layer informs the database, that the write was completed but actually not. In order to use this detection feature physical Standby has to be used and DB_LOST_WRITE_PROTECT parameter must be set "TYPICAL" on both Primary and standby. When lost-write is detected, standby recovery stops and the way to get rid of this corruption is to failover to standby database.

RMAN Integration

Rman integration is not a new feature. The ability to use a backup that is taken on the Physical standby for primary and vice versa exists both in 10g and 11g. You must use a Recovery Catalog in order to take advantage of RMAN&Data Guard integration. And beginning with 11g, for "Block Change Tracking" feature of RMAN, which records the changed blocks for incremental backups, standby databases can be used.

Integration with Oracle Applications

Business Intelligence:

- Directs write operations to primary
- All read operations to active data guard standby

Toplink:

- Applications developed with Oracle TopLink are able to be configured as "Active Data Guard aware"

Siebel CRM:

- An ongoing study,
- Write operations will work on primary
- Read operations will work on standby
- Automatic direction to primary in a case of lag

Contact address:

Emre Baransel
Turkcell
Turkcell Plaza Eskisehir Yolu NO:264
06450 Ankara-Turkey

Phone: +9(0)5322104531
Email ebaransel@yahoo.com
Internet: www.emrebaransel.tel